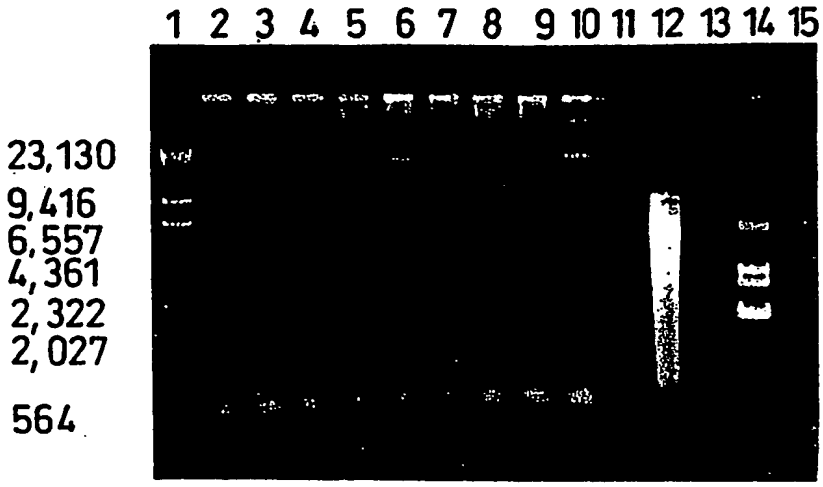


Fig. 1

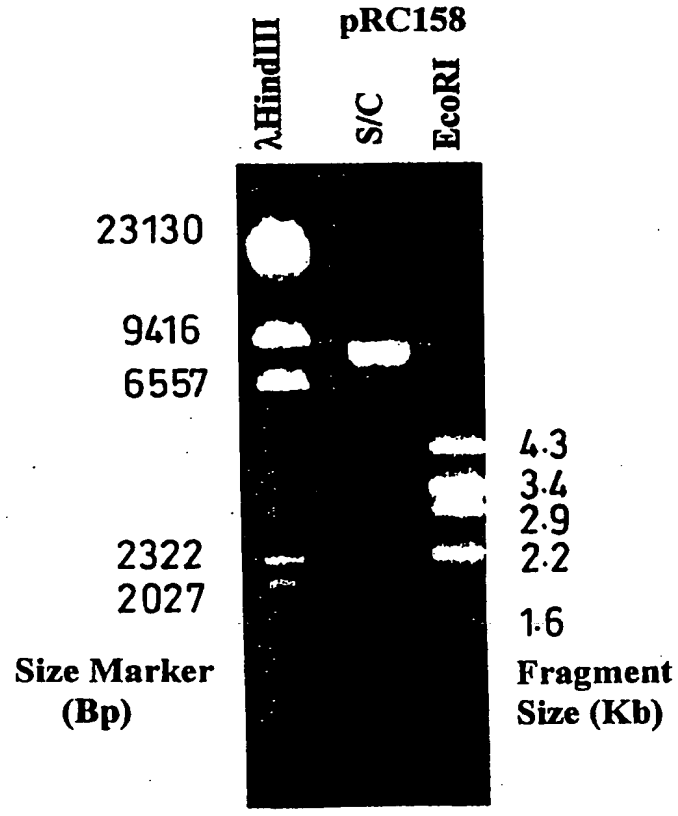
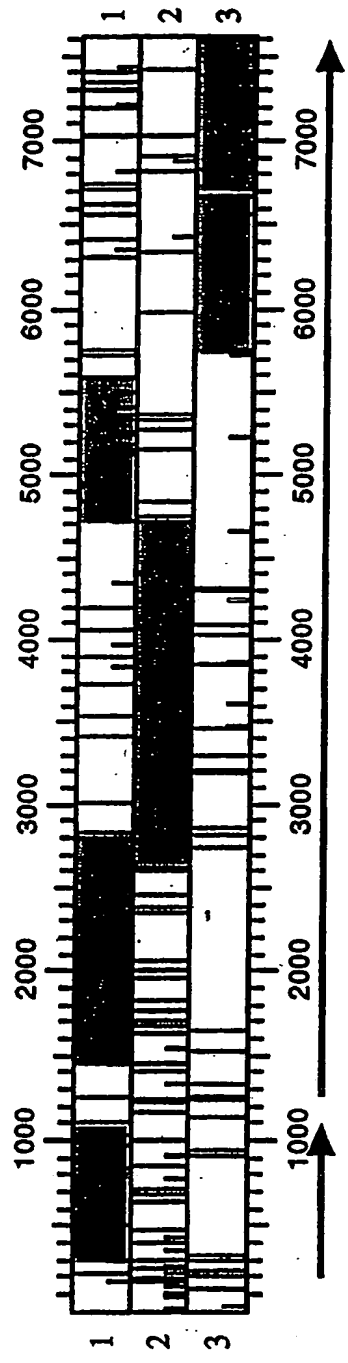


Fig. 2

Fig 3



Gene	Initiator Codon	Terminator Codon	Molecular Weight
Regulator	295	1035	27102
Transport	1450	2805	47433
Monooxygenase	2810	4720	69650
Hydroxymuconic semialdehyde hydrolase	4717	5586	32770
Catechol 2, 3-dioxygenase	5721	6665	33894
Alcohol dehydrogenase	6711	7580	30586

Fig. 4A

```
        10                  30                  50
GAATTCCATGTTCTTCTCCTTGCATGTGGCCCGCGTTGCCGAGGGCACTGCTCGGCCTGT
CTTAAGGTACAAGAAGAGGAACGTACACCGGGCGCAACGGCTCCCGTGACGAGCCGGACA

        70                  90                  110
CGCCCGCAGAGGGCGCATGTCCGGGTGCCTGGATATGGCGCGTACGGCGTGCCCTCCGGC
GCGGGCGTCTCCCGCGTACAGGCCCACGGACCTATACCGCGCATGCCGCACGGGAGGCCG

        130                 150                 170
GTTAACCCCGAGGTTGGCCACGATGCCCCGGCCATCAGGTCTGGAATGCTAGCGTTCCAG
CAATTGGGGCTCCAACCGGTGCTACGGGGCCGGTAGTCCAGACCTTACGATCGCAAGGTC

        190                 210                 230
ACGAAGGTAACCCACAGTGACTCACACCACAAGTACTAGAATGCAAGCTGTTGCGGTGAG
TGCTTCCATTGGGTGTCACTGAGTGTGGTGTTCATGATCTTACGTTCGACAACGCCACTC

        250                 270                 290
CGCCGCGGCATAAGGGGGAGCCATGTCCGGGACGCCGACGGAAAGCCTGACTCGATGACC
GCGGCGCCGTATTCCCCCTCGGTACAGGCCCTGCGGCTGCCTTTCGGACTGAGCTACTGG
                                                    M   T

        310                 330                 350
ACCACCGACACCGGCCCCAAGCCGGGCAGTGAGGCCGCCGCCCTGCTCGCCAATGTCCGC
TGGTGGCTGTGGCCGGGGTTCGGCCCGTCACTCCGGCGGCGGGACGAGCGGTTACAGGCG
T   T   D   T   G   P   K   P   G   S   E   A   A   A   L   L   A   N   V   R

        370                 390                 410
ACCTCGGGGGCGCGGCTGTCCTCCGCGTTGTACGACATTCTGAAGAACCGGCTGCTCGAA
TGGAGCCCCCGCGCCGACAGGAGGCGCAACATGCTGTAAGACTTCTTGGCCGACGAGCTT
T   S   G   A   R   L   S   S   A   L   Y   D   I   L   K   N   R   L   L   E

        430                 450                 470
GGGCGCTATGCGGCAGGCGAGAAGATCGTCGTCGAGTCGATCCGGCAAGAGTTCGGGGTG
CCCGCGATACGCCGTCCGCTCTTCTAGCAGCAGCTCAGCTAGGCCGTTCTCAAGCCCCAC
G   R   Y   A   A   G   E   K   I   V   V   E   S   I   R   Q   E   F   G   V

        490                 510                 530
AGCAAGCAGCCCGTCATGGACGCTCTGCGCCGCCTGTCCAGCGACAAGCTGGTCCACATC
TCGTTCGTCGGGCAGTACCTGCGAGACGCGGCGGACAGGTCGCTGTTCGACCAGGTGTAG
S   K   Q   P   V   M   D   A   L   R   R   L   S   S   D   K   L   V   H   I

        550                 570                 590
GTTCCCCAGGTCGGTTGCGAGGTCGTCTCCTACGCCCCGCGCGAAGTGGAAGACTTCTAC
CAAGGGGTCCAGCCAACGCTCCAGCAGAGGATGCGGGGCGCGCTTCACCTTCTGAAGATG
V   P   Q   V   G   C   E   V   V   S   Y   A   P   R   E   V   E   D   F   Y

        610                 630                 650
ACCCTGTTCGGCGGTTTCGAAGGGACCATCGCCGCGGTAGCGGCCTCCCGGCGGACCGAG
```

Fig. 4B

```
TGGGACAAGCCGCCAAAGCTTCCCTGGTAGCGGCGCCATCGCCGGAGGGCCGCCTGGCTC
T  L  F  G  G  F  E  G  T  I  A  A  V  A  A  S  R  R  T  E

      670                 690                 710

GCCCAGTTGCTGGAGCTGGACCTGATCTCGGCGCGGGTCGACGCCCTGATCACCTCCCAC
CGGGTCAACGACCTCGACCTGGACTAGAGCCGCGCCCAGCTGCGGGACTAGTGGAGGGTG
A  Q  L  L  E  L  D  L  I  S  A  R  V  D  A  L  I  T  S  H

      730                 750                 770

GACCCGGTGGTCCGCGCCCGCGGGTACCGCGTGCACAACCGGGAGTTCCATGCGGCCATC
CTGGGCCACCAGGCGCGGGCGCCCATGGCGCACGTGTTGGCCCTCAAGGTACGCCGGTAG
D  P  V  V  R  A  R  G  Y  R  V  H  N  R  E  F  H  A  A  I

      790                 810                 830

CACGCGATGGCGCACTCGCGGATCATGGAGGAGACCAGCCAGCGAATGTGGGATCTGTCG
GTGCGCTACCGCGTGAGCGCCTAGTACCTCCTCTGGTCGGTCGCTTACACCCTAGACAGC
H  A  M  A  H  S  R  I  M  E  E  T  S  Q  R  M  W  D  L  S

      850                 870                 890

GACTTCTTGATCAACACCACCGGCATCACCAACCCGCTCTCGAGCGCACTGCCCGACCGG
CTGAAGAACTAGTTGTGGTGGCCGTAGTGGTTGGGCGAGAGCTCGCGTGACGGGCTGGCC
D  F  L  I  N  T  T  G  I  T  N  P  L  S  S  A  L  P  D  R

      910                 930                 950

CAGCATGACCACCACGAAATCACCGAGGCCATCCGCAACCGTGACGCAGCTGCCGCCCGC
GTCGTACTGGTGGTGCTTTAGTGGCTCCGGTAGGCGTTGGCACTGCGTCGACGGCGGGCG
Q  H  D  H  H  E  I  T  E  A  I  R  N  R  D  A  A  A  A  R

      970                 990                 1010

GAGGCCATGGAACGCCACATCGTCGGCACCATCGCAGTAATCCGCGACGAATCCAACGCC
CTCCGGTACCTTGCGGTGTAGCAGCCGTGGTAGCGTCATTAGGCGCTGCTTAGGTTGCGG
E  A  M  E  R  H  I  V  G  T  I  A  V  I  R  D  E  S  N  A

      1030                1050                1070

CAGCTGCCGAGCTAGACCCCGATACCCGGGCCATCGACCGGCTCCGCTATCGCGCCACCT
GTCGACGGCTCGATCTGGGGCTATGGGCCCGGTAGCTGGCCGAGGCGATAGCGCGGTGGA
Q  L  P  S  *

      1090                1110                1130

ACGCCGAGGGGGGACTCTCGGCCGTAGCGCTGCAGACGATCCACCGGCACCCTCCACGCT
TGCGGCTCCCCCCTGAGAGCCGGCATCGCGACGTCTGCTAGGTGGCCGTGGGAGGTGCGA

      1150                1170                1190

GACCCCTGTCTCGCCCTAGAGGGCCGGCGCGCCGTCGATCACCTTTACCCTCATCCAGAG
CTGGGGACAGAGCGGGATCTCCCGGCCGCGCGGCAGCTAGTGGAAATGGGAGTAGGTCTC

      1210                1230                1250

ACTTGCGTCACCCTCTATGCCCGAGTAGCGTCTGAACTAGACGTCTAGCATTCTAGTTGA
TGAACGCAGTGGGAGATACGGGCTCATCGCAGACTTGATCTGCAGATCGTAAGATCAACT

      1270                1290                1310

GTGCTCCCTCTCGAAGATTCTCCAGAGAACCCCTCTCGAACATCCCCAGAAGAAAGGAGC
```

Fig. 4C

```
CACGAGGGAGAGCTTCTAAGAGGTCTCTTGGGGAGAGCTTGTAGGGGTCTTCTTTCCTCG

      1330                 1350                 1370
      .                    .                    .
GGCCATGACGACCGCTTCGCACGCATCGTCCTTCGGGGCACGAGCCCACTTCCGCCCACA
CCGGTACTGCTGGCGAAGCGTGCGTAGCAGGAAGCCCCGTGCTCGGGTGAAGGCGGGTGT

      1390                 1410                 1430
      .                    .                    .
GATCGGGGAAGCCCGACCGTGAGCACCACACCTACCTCCCCGACGAAGACCTCACCGCTG
CTAGCCCCTTCGGGCTGGCACTCGTGGTGTGGATGGAGGGGCTGCTTCTGGAGTGGCGAC

      1450                 1470                 1490
      .                    .                    .
CGGGTAGCGATGGCCAGCTTCATCGGTACCACCGTCGAGTACTACGACTTCTTCATCTAC
GCCCATCGCTACCGGTCGAAGTAGCCATGGTGGCAGCTCATGATGCTGAAGAAGTAGATG
         M  A  S  F  I  G  T  T  V  E  Y  Y  D  F  F  I  Y

      1510                 1530                 1550
      .                    .                    .
GGCACCGCGGCCGCGCTGGTATTCCCTGAGTTGTTCTTCCCGGATGTCTCGTCCGCGATC
CCGTGGCGCCGGCGCGACCATAAGGGACTCAACAAGAAGGGCCTACAGAGCAGGCGCTAG
G  T  A  A  A  L  V  F  P  E  L  F  F  P  D  V  S  S  A  I

      1570                 1590                 1610
      .                    .                    .
GGAATCCTGTTGTCGTTCGCGACCTTCAGCGTTGGGTTCCTCGCCCGCCCGCTGGGTGGC
CCTTAGGACAACAGCAAGCGCTGGAAGTCGCAACCCAAGGAGCGGGCGGGCGACCCACCG
G  I  L  L  S  F  A  T  F  S  V  G  F  L  A  R  P  L  G  G

      1630                 1650                 1670
      .                    .                    .
ATAGTGTTCGGGCACTTCGGTGACCGGGTCGGCCGCAAGCAGATGCTGGTGATCTCCCTG
TATCACAAGCCCGTGAAGCCACTGGCCCAGCCGGCGTTCGTCTACGACCACTAGAGGGAC
I  V  F  G  H  F  G  D  R  V  G  R  K  Q  M  L  V  I  S  L

      1690                 1710                 1730
      .                    .                    .
GTCGGAATGGGCTCGGCCACCGTACTGATGGGATTGTTGCCCGGTTACGCCCAAATCGGG
CAGCCTTACCCGAGCCGGTGGCATGACTACCCTAACAACGGGCCAATGCGGGTTTAGCCC
V  G  M  G  S  A  T  V  L  M  G  L  L  P  G  Y  A  Q  I  G

      1750                 1770                 1790
      .                    .                    .
ATCGCCGCCCCCATCCTGCTGACCCTGCTGCGCCTGGTGCAGGGCTTTGCCGTCGGCGGC
TAGCGGCGGGGGTAGGACGACTGGGACGACGCGGACCACGTCCCGAAACGGCAGCCGCCG
I  A  A  P  I  L  L  T  L  L  R  L  V  Q  G  F  A  V  G  G

      1810                 1830                 1850
      .                    .                    .
GAGTGGGGTGGAGCCACCCTGATGGCCGTCGAGCACGCCCCCACCGCGAAGAAGGGCTTT
CTCACCCCACCTCGGTGGGACTACCGGCAGCTCGTGCGGGGGTGGCGCTTCTTCCCGAAA
E  W  G  G  A  T  L  M  A  V  E  H  A  P  T  A  K  K  G  F

      1870                 1890                 1910
      .                    .                    .
TTCGGATCCTTCTCCCAGATGGGGGCACCCGCCGGGACCAGCGTCGCAACCCTGGCGTTC
AAGCCTAGGAAGAGGGTCTACCCCCGTGGGCGGCCCTGGTCGCAGCGTTGGGACCGCAAG
F  G  S  F  S  Q  M  G  A  P  A  G  T  S  V  A  T  L  A  F

      1930                 1950                 1970
      .                    .                    .
TTCGCGGTCTCCCAATTGCCCGACGAGCAGTTCCTGAGTTGGGGCTGGCGACTGCCGTTC
```

Fig. 4D

```
AAGCGCCAGAGGGTTAACGGGCTGCTCGTCAAGGACTCAACCCCGACCGCTGACGGCAAG
F  A  V  S  Q  L  P  D  E  Q  F  L  S  W  G  W  R  L  P  F

      1990                2010                2030
      .                   .                   .
CTGTTCAGCGCGGTGCTGATCGTGATCGGGCTGTTCATTCGCCTGTCCCTGGCCGAAAGC
GACAAGTCGCGCCACGACTAGCACTAGCCCGACAAGTAAGCGGACAGGGACCGGCTTTCG
L  F  S  A  V  L  I  V  I  G  L  F  I  R  L  S  L  A  E  S

      2050                2070                2090
      .                   .                   .
CCCGACTTCGCCGAGGTGAAGGCACAGAGCGCCGTGGTGCGAATGCCGATCGCCGAAGCG
GGGCTGAAGCGGCTCCACTTCCGTGTCTCGCGGCACCACGCTTACGGCTAGCGGCTTCGC
P  D  F  A  E  V  K  A  Q  S  A  V  V  R  M  P  I  A  E  A

      2110                2130                2150
      .                   .                   .
TTCCGCAAGCACTGGAAGGAAATTCTCCTCATCGCGGGCACCTACCTGTCCCAAGGAGTG
AAGGCGTTCGTGACCTTCCTTTAAGAGGAGTAGCGCCCGTGGATGGACAGGGTTCCTCAC
F  R  K  H  W  K  E  I  L  L  I  A  G  T  Y  L  S  Q  G  V

      2170                2190                2210
      .                   .                   .
TTCGCCTATATCTGCATGGCCTACCTCGTCTCCTACGGCACCACCGTCGCGGGGATCAGC
AAGCGGATATAGACGTACCGGATGGAGCAGAGGATGCCGTGGTGGCAGCGCCCCTAGTCG
F  A  Y  I  C  M  A  Y  L  V  S  Y  G  T  T  V  A  G  I  S

      2230                2250                2270
      .                   .                   .
CGCACCTTCGCCCTGGCCGGAGTATTCGTCGCCGGCATCGTCGCCGTCCTCCTCTACCTC
GCGTGGAAGCGGGACCGGCCTCATAAGCAGCGGCCGTAGCAGCGGCAGGAGGAGATGGAG
R  T  F  A  L  A  G  V  F  V  A  G  I  V  A  V  L  L  Y  L

      2290                2310                2330
      .                   .                   .
GTGTTCGGCGCTCTGTCCGACACTTTCGGCCGCAAGACCATGTACCTGCTCGGCGCCGCC
CACAAGCCGCGAGACAGGCTGTGAAAGCCGGCGTTCTGGTACATGGACGAGCCGCGGCGG
V  F  G  A  L  S  D  T  F  G  R  K  T  M  Y  L  L  G  A  A

      2350                2370                2390
      .                   .                   .
GCGATGGGTGTGGTGATCGCCCCCGCCTTCGCACTGATCAACACCGGCAACCCGTGGCTG
CGCTACCCACACCACTAGCGGGGGCGGAAGCGTGACTAGTTGTGGCCGTTGGGCACCGAC
A  M  G  V  V  I  A  P  A  F  A  L  I  N  T  G  N  P  W  L

      2410                2430                2450
      .                   .                   .
TTCATGGCCGCGCAGGTGCTGGTCTTCGGAATTGCAATGGCCCCCGCCGCCGGCGTGACA
AAGTACCGGCGCGTCCACGACCAGAAGCCTTAACGTTACCGGGGGCGGCGGCCGCACTGT
F  M  A  A  Q  V  L  V  F  G  I  A  M  A  P  A  A  G  V  T

      2470                2490                2510
      .                   .                   .
GGCTCCCTGTTCACGATGGTCTTCGACGCGGACGTGCGCTACAGCGGTGTCTCTATCGGC
CCGAGGGACAAGTGCTACCAGAAGCTGCGCCTGCACGCGATGTCGCCACAGAGATAGCCG
G  S  L  F  T  M  V  F  D  A  D  V  R  Y  S  G  V  S  I  G

      2530                2550                2570
      .                   .                   .
TACACCATCTCCCAGGTCGCCGGCTCCGCGTTCGCCCCGACGATCGCGACCGCCTTGTAC
ATGTGGTAGAGGGTCCAGCGGCCGAGGCGCAAGCGGGGCTGCTAGCGCTGGCGGAACATG
Y  T  I  S  Q  V  A  G  S  A  F  A  P  T  I  A  T  A  L  Y
```

Fig. 4E

```
      2590                2610                2630
GCCTCCACCAACACCAGCAACTCGATCGTGACCTACCTGCTGATCGTCTCGGCCATCTCG
CGGAGGTGGTTGTGGTCGTTGAGCTAGCACTGGATGGACGACTAGCAGAGCCGGTAGAGC
A  S  T  N  T  S  N  S  I  V  T  Y  L  L  I  V  S  A  I  S

      2650                2670                2690
ATCGTCTCGGTGATCCTGCTGCCCGGCGGCTGGGGGCGCAAGGGCGCTGCGAGCCAGCTC
TAGCAGAGCCACTAGGACGACGGGCCGCCGACCCCCGCGTTCCCGCGACGCTCGGTCGAG
I  V  S  V  I  L  L  P  G  G  W  G  R  K  G  A  A  S  Q  L

      2710                2730                2750
ACTCGCGACCAGGCCACCTCCACACCGAAAATGCCTGACACCGAAACATTTTCGACTCGG
TGAGCGCTGGTCCGGTGGAGGTGTGGCTTTTACGGACTGTGGCTTTGTAAAAGCTGAGCC
T  R  D  Q  A  T  S  T  P  K  M  P  D  T  E  T  F  S  T  R

      2770                2790                2810
ACAGTTCCGGACACCGCAGCATCCCTGCGCGTCCTCGACAAGTGAAGTGATGACAGACAT
TGTCAAGGCCTGTGGCGTCGTAGGGACGCGCAGGAGCTGTTCACTTCACTACTGTCTGTA
T  V  P  D  T  A  A  S  L  R  V  L  D  K  *     M  T  D  M

      2830                2850                2870
GAGTGACCACGACCGCACCTCCTACGACACCGACGTCGTGATCGTCGGCCTCGGCCCCGC
CTCACTGGTGCTGGCGTGGAGGATGCTGTGGCTGCAGCACTAGCAGCCGGAGCCGGGGCG
 S  D  H  D  R  T  S  Y  D  T  D  V  V  I  V  G  L  G  P  A

      2890                2910                2930
CGGTGGCACAGCGGCGCTTGCCCTGGCCAGCTACGGCATCCGCGTTCACGCCGTCTCGAT
GCCACCGTGTCGCCGCGAACGGGACCGGTCGATGCCGTAGGCGCAAGTGCGGCAGAGCTA
 G  G  T  A  A  L  A  L  A  S  Y  G  I  R  V  H  A  V  S  M

      2950                2970                2990
GTTCCCCTGGGTGGCGAACTCGCCGCGCGCGCACATCACCAACCAGCGCGCCGTCGAAGT
CAAGGGGACCCACCGCTTGAGCGGCGCGCGCGTGTAGTGGTTGGTCGCGCGGCAGCTTCA
 F  P  W  V  A  N  S  P  R  A  H  I  T  N  Q  R  A  V  E  V

      3010                3030                3050
GCTGCGTGACCTGGGCGTCGAAGACGAGGCGCGCAACTACGCCACCCCGTGGGACCAGAT
CGACGCACTGGACCCGCAGCTTCTGCTCCGCGCGTTGATGCGGTGGGGCACCCTGGTCTA
 L  R  D  L  G  V  E  D  E  A  R  N  Y  A  T  P  W  D  Q  M

      3070                3090                3110
GGGCGACACGCTGTTCACCACGAGCCTGGCCGGCGAGGAGATCGTCCGGATGCAGACCTG
CCCGCTGTGCGACAAGTGGTGCTCGGACCGGCCGCTCCTCTAGCAGGCCTACGTCTGGAC
 G  D  T  L  F  T  T  S  L  A  G  E  E  I  V  R  M  Q  T  W

      3130                3150                3170
GGGTACGGGCGATATCCGCTACGGGGACTACCTGTCCGGAAGCCCCTGCACGATGCTCGA
CCCATGCCCGCTATAGGCGATGCCCCTGATGGACAGGCCTTCGGGGACGTGCTACGAGCT
 G  T  G  D  I  R  Y  G  D  Y  L  S  G  S  P  C  T  M  L  D

      3190                3210                3230
CATTCCGCAGCCCCTGATGGAGCCGGTGCTGATCAAGAACGCCGCCGAACGTGGTGCGGT
```

Fig. 4F

```
GTAAGGCGTCGGGGACTACCTCGGCCACGACTAGTTCTTGCGGCGGCTTGCACCACGCCA
 I  P  Q  P  L  M  E  P  V  L  I  K  N  A  A  E  R  G  A  V

      3250                3270                3290

CATCAGCTTCAACACCGAATACCTCGACCACGCCCAGGACGAGGACGGGGTGACCGTCCG
GTAGTCGAAGTTGTGGCTTATGGAGCTGGTGCGGGTCCTGCTCCTGCCCCACTGGCAGGC
 I  S  F  N  T  E  Y  L  D  H  A  Q  D  E  D  G  V  T  V  R

      3310                3330                3350

GTTCCGCGACGTCCGCTCGGGCACCGTGTTCACCCAGCGAGCCCGCTTCCTGCTCGGTTT
CAAGGCGCTGCAGGCGAGCCCGTGGCACAAGTGGGTCGCTCGGGCGAAGGACGAGCCAAA
 F  R  D  V  R  S  G  T  V  F  T  Q  R  A  R  F  L  L  G  F

      3370                3390                3410

CGACGGCGCACGATCGAAGATCGCCGAACAGATCGGGCTTCCGTTCGAAGGTGAACTCGC
GCTGCCGCGTGCTAGCTTCTAGCGGCTTGTCTAGCCCGAAGGCAAGCTTCCACTTGAGCG
 D  G  A  R  S  K  I  A  E  Q  I  G  L  P  F  E  G  E  L  A

      3430                3450                3470

CCGCGCCGGTACCGCGTACATCCTGTTCAACGCGGACCTGAGCAAATATGTCGCTCATCG
GGCGCGGCCATGGCGCATGTAGGACAAGTTGCGCCTGGACTCGTTTATACAGCGAGTAGC
 R  A  G  T  A  Y  I  L  F  N  A  D  L  S  K  Y  V  A  H  R

      3490                3510                3530

GCCGAGCATCTTGCACTGGATCGTCAACTCGAAGGCCGGTTTCGGTGAGATCGGCATGGG
CGGCTCGTAGAACGTGACCTAGCAGTTGAGCTTCCGGCCAAAGCCACTCTAGCCGTACCC
 P  S  I  L  H  W  I  V  N  S  K  A  G  F  G  E  I  G  M  G

      3550                3570                3590

TCTGCTGCGCGCGATCCGACCGTGGGACCAGTGGATCGCCGGCTGGGGCTTCGACATGGC
AGACGACGCGCGCTAGGCTGGCACCCTGGTCACCTAGCGGCCGACCCCGAAGCTGTACCG
 L  L  R  A  I  R  P  W  D  Q  W  I  A  G  W  G  F  D  M  A

      3610                3630                3650

GAACGGCGAGCCGGATGTCTCCGACGACGTTGTCCTCGAACAGATCCGGACCCTCGTCGG
CTTGCCGCTCGGCCTACAGAGGCTGCTGCAACAGGAGCTTGTCTAGGCCTGGGAGCAGCC
 N  G  E  P  D  V  S  D  D  V  V  L  E  Q  I  R  T  L  V  G

      3670                3690                3710

CGACCCGCACCTGGACGTCGAGATCGTGTCGAGGTCCTTCTGGTACGTCAACCGGCAGTG
GCTGGGCGTGGACCTGCAGCTCTAGCACAGCTCCAGGAAGACCATGCAGTTGGCCGTCAC
 D  P  H  L  D  V  E  I  V  S  R  S  F  W  Y  V  N  R  Q  W

      3730                3750                3770

GGCTGAGCACTACCAGTCCGGTCGAGTGTTCTGCGGCGGCGACGCGGTGCACCGGCATCC
CCGACTCGTGATGGTCAGGCCAGCTCACAAGACGCCGCCGCTGCGCCACGTGGCCGTAGG
 A  E  H  Y  Q  S  G  R  V  F  C  G  G  D  A  V  H  R  H  P

      3790                3810                3830

GCCGAGCAGCGGGCTGGGCTCGAACACGTCCATGCAGGACGCGTTCAACCTGGCATGGAA
CGGCTCGTCGCCCGACCCGAGCTTGTGCAGGTACGTCCTGCGCAAGTTGGACCGTACCTT
 P  S  S  G  L  G  S  N  T  S  M  Q  D  A  F  N  L  A  W  K
```

Fig. 4G

```
      3850                3870                3890
GATCGCGTTCGTCGTGAAGGGGTATGCAGGACCGGGTCTGCTCGAGTCCTACTCTCCTGA
CTAGCGCAAGCAGCACTTCCCCATACGTCCTGGCCCAGACGAGCTCAGGATGAGAGGACT
 I  A  F  V  V  K  G  Y  A  G  P  G  L  L  E  S  Y  S  P  E

      3910                3930                3950
GCGTGTTCCGGTCGGCAAACAGATCGTCGCTCGCGCCAACCAGTCCCGCAAGGACTACGC
CGCACAAGGCCAGCCGTTTGTCTAGCAGCGAGCGCGGTTGGTCAGGGCGTTCCTGATGCG
 R  V  P  V  G  K  Q  I  V  A  R  A  N  Q  S  R  K  D  Y  A

      3970                3990                4010
CGGGCTGCGCGAATGGTTCGATCACGAGAGCGACGACCCGGTCGCCGCCGGCCTGGCAAA
GCCCGACGCGCTTACCAAGCTAGTGCTCTCGCTGCTGGGCCAGCGGCGGCCGGACCGTTT
 G  L  R  E  W  F  D  H  E  S  D  D  P  V  A  A  G  L  A  K

      4030                4050                4070
GTTGAAGGAACCCTCGTCCGAAGGTGTTGCTCTGCGTGAGCGGCTGTACGAGGCGCTGGA
CAACTTCCTTGGGAGCAGGCTTCCACAACGAGACGCACTCGCCGACATGCTCCGCGACCT
 L  K  E  P  S  S  E  G  V  A  L  R  E  R  L  Y  E  A  L  E

      4090                4110                4130
GGTGAAGAACGCCGAATTCAACGCCCAGGGCGTCGAACTCAACCAGCGCTACACCTCGTC
CCACTTCTTGCGGCTTAAGTTGCGGGTCCCGCAGCTTGAGTTGGTCGCGATGTGGAGCAG
 V  K  N  A  E  F  N  A  Q  G  V  E  L  N  Q  R  Y  T  S  S

      4150                4170                4190
CGCGGTCGTTCCCGACCCCGAGGCGGGCGAGGAAGTGTGGGTGCGCGATCGTGAGCTGTA
GCGCCAGCAAGGGCTGGGGCTCCGCCCGCTCCTTCACACCCACGCGCTAGCACTCGACAT
 A  V  V  P  D  P  E  A  G  E  E  V  W  V  R  D  R  E  L  Y

      4210                4230                4250
CCTGCAGGCCACCACCCGGCCGGGCGCGAAGCTGCCGCATGCGTGGCTGGTCGGCGCCGA
GGACGTCCGGTGGTGGGCCGGCCCGCGCTTCGACGGCGTACGCACCGACCAGCCGCGGCT
 L  Q  A  T  T  R  P  G  A  K  L  P  H  A  W  L  V  G  A  D

      4270                4290                4310
CGGAACCCGCATCTCCACCCTCGACGTCACCGGCAAGGGAATGATGACCCTGCTGACCGG
GCCTTGGGCGTAGAGGTGGGAGCTGCAGTGGCCGTTCCCTTACTACTGGGACGACTGGCC
 G  T  R  I  S  T  L  D  V  T  G  K  G  M  M  T  L  L  T  G

      4330                4350                4370
ACTCGGCGGCCAGGCATGGAAGCGTGCCGCCGCCAAACTCGACCTGCCGTTCCTGCGGAC
TGAGCCGCCGGTCCGTACCTTCGCACGGCGGCGGTTTGAGCTGGACGGCAAGGACGCCTG
 L  G  G  Q  A  W  K  R  A  A  A  K  L  D  L  P  F  L  R  T

      4390                4410                4430
CGTCGTTGTCGGCGAACCCGGCACCATCGACCCTTACGGATACTGGCGGCGGGTCCGCGA
GCAGCAACAGCCGCTTGGGCCGTGGTAGCTGGGAATGCCTATGACCGCCGCCCAGGCGCT
 V  V  V  G  E  P  G  T  I  D  P  Y  G  Y  W  R  R  V  R  D

      4450                4470                4490
CATCGACGAGGCCGGCGCCCTGCTCGTGCGGCCCGACGGCTACGTCGCGTGGCGACACAG
```

Fig. 4H

```
GTAGCTGCTCCGGCCGCGGGACGAGCACGCCGGGCTGCCGATGCAGCGCACCGCTGTGTC
 I  D  E  A  G  A  L  L  V  R  P  D  G  Y  V  A  W  R  H  S

      4510                 4530                 4550

TGCTCCGGTCTGGGACGACACCGAAGCGCTCACCAGCCTCGAGAACGCTCTCACCGCGGT
ACGAGGCCAGACCCTGCTGTGGCTTCGCGAGTGGTCGGAGCTCTTGCGAGAGTGGCGCCA
 A  P  V  W  D  D  T  E  A  L  T  S  L  E  N  A  L  T  A  V

      4570                 4590                 4610

CCTCGACCACTCGGCCAGCGACAACGGGAACCCGAGCGGCACAAACGAGCCGCAGTACAG
GGAGCTGGTGAGCCGGTCGCTGTTGCCCTTGGGCTCGCCGTGTTTGCTCGGCGTCATGTC
 L  D  H  S  A  S  D  N  G  N  P  S  G  T  N  E  P  Q  Y  S

      4630                 4650                 4670

CACCCGGGCCGTGCCGATCGTCGTTCCGCACGTTACCGCCGAGGATGCAGCACCAGCTTC
GTGGGCCCGGCACGGCTAGCAGCAAGGCGTGCAATGGCGGCTCCTACGTCGTGGTCGAAG
 T  R  A  V  P  I  V  V  P  H  V  T  A  E  D  A  A  P  A  S

      4690                 4710                 4730

CGCCACCCGCACCACCACAGTCGAGGGAGAGAACCGATGACCCGTCCTTACACCAGCGTC
GCGGTGGGCGTGGTGGTGTCAGCTCCCTCTCTTGGCTACTGGGCAGGAATGTGGTCGCAG
 A  T  R  T  T  T  V  E  G  E  N  R  *
                                    M  T  R  P  Y  T  S  V

      4750                 4770                 4790

TGGGACGACCTGAACCAGGTCGAGTTCAGCCAGGGATTCATCCAGGCCGGCCCCTACCGG
ACCCTGCTGGACTTGGTCCAGCTCAAGTCGGTCCCTAAGTAGGTCCGGCCGGGGATGGCC
 W  D  D  L  N  Q  V  E  F  S  Q  G  F  I  Q  A  G  P  Y  R

      4810                 4830                 4850

ACCCGATACCTGCACGCCGGCGATTCGTCCAAGCCCACGCTGATCCTGCTGCACGGCATC
TGGGCTATGGACGTGCGGCCGCTAAGCAGGTTCGGGTGCGACTAGGACGACGTGCCGTAG
 T  R  Y  L  H  A  G  D  S  S  K  P  T  L  I  L  L  H  G  I

      4870                 4890                 4910

ACCGGCCACGCCGAGGCGTACGTGCGCAATCTGCGCTCGCATTCCGAGCACTTCAACGTC
TGGCCGGTGCGGCTCCGCATGCACGCGTTAGACGCGAGCGTAAGGCTCGTGAAGTTGCAG
 T  G  H  A  E  A  Y  V  R  N  L  R  S  H  S  E  H  F  N  V

      4930                 4950                 4970

TGGGCAATCGACTTCATCGGCCACGGCTATTCGACCAAGCCCGACCACCCGCTCGAGATC
ACCCGTTAGCTGAAGTAGCCGGTGCCGATAAGCTGGTTCGGGCTGGTGGGCGAGCTCTAG
 W  A  I  D  F  I  G  H  G  Y  S  T  K  P  D  H  P  L  E  I

      4990                 5010                 5030

AAGCACTACATCGACCACGTGCTGCAGTTGCTGGACGCCATCGGCGTCGAGAAGGCCTCG
TTCGTGATGTAGCTGGTGCACGACGTCAACGACCTGCGGTAGCCGCAGCTCTTCCGGAGC
 K  H  Y  I  D  H  V  L  Q  L  L  D  A  I  G  V  E  K  A  S

      5050                 5070                 5090

TTTTCCGGGGAGTCTCTCGGCGGTTGGGTCACCGCCCAGTTCGCGCACGACCATCCCGAG
AAAAGGCCCCTCAGAGAGCCGCCAACCCAGTGGCGGGTCAAGCGCGTGCTGGTAGGGCTC
 F  S  G  E  S  L  G  G  W  V  T  A  Q  F  A  H  D  H  P  E
```

Fig. 4I

```
      5110                 5130                 5150
AAGGTCGACCGGATCGTGCTCAACACCATGGGCGGCACCATGGCCAACCCTCAGGTGATG
TTCCAGCTGGCCTAGCACGAGTTGTGGTACCCGCCGTGGTACCGGTTGGGAGTCCACTAC
K  V  D  R  I  V  L  N  T  M  G  G  T  M  A  N  P  Q  V  M

      5170                 5190                 5210
GAACGTCTCTATACCCTGTCGATGGAAGCGGCGAAGGACCCGAGCTGGGAACGCGTCAAA
CTTGCAGAGATATGGGACAGCTACCTTCGCCGCTTCCTGGGCTCGACCCTTGCGCAGTTT
E  R  L  Y  T  L  S  M  E  A  A  K  D  P  S  W  E  R  V  K

      5230                 5250                 5270
GCACGCCTCGAATGGCTCATGGCCGACCCGACCATGGTCACCGACGACCTGATCCGCACC
CGTGCGGAGCTTACCGAGTACCGGCTGGGCTGGTACCAGTGGCTGCTGGACTAGGCGTGG
A  R  L  E  W  L  M  A  D  P  T  M  V  T  D  D  L  I  R  T

      5290                 5310                 5330
CGCCAGGCCATCTTCCAGCAGCCGGATTGGCTCAAGGCCTGCGAGATGAACATGGCACTG
GCGGTCCGGTAGAAGGTCGTCGGCCTAACCGAGTTCCGGACGCTCTACTTGTACCGTGAC
R  Q  A  I  F  Q  Q  P  D  W  L  K  A  C  E  M  N  M  A  L

      5350                 5370                 5390
CAGGACCTCGAAACCCGCAAGCGGAACATGATCACCGACGCCACTCTCAACGGCATCACG
GTCCTGGAGCTTTGGGCGTTCGCCTTGTACTAGTGGCTGCGGTGAGAGTTGCCGTAGTGC
Q  D  L  E  T  R  K  R  N  M  I  T  D  A  T  L  N  G  I  T

      5410                 5430                 5450
GTGCCCGCGATGGTGCTGTGGACCACCAAGGACCCCTCCGGTCCGGTCGACGAAGCCAAG
CACGGGCGCTACCACGACACCTGGTGGTTCCTGGGGAGGCCAGGCCAGCTGCTTCGGTTC
V  P  A  M  V  L  W  T  T  K  D  P  S  G  P  V  D  E  A  K

      5470                 5490                 5510
CGCATCGCCTCCCACATCCCGGGCGCCAAGCTGGCCATCATGGAGAACTGTGGCCACTGG
GCGTAGCGGAGGGTGTAGGGCCCGCGGTTCGACCGGTAGTACCTCTTGACACCGGTGACC
R  I  A  S  H  I  P  G  A  K  L  A  I  M  E  N  C  G  H  W

      5530                 5550                 5570
CCCCAGTACGAGGACCCCGAGACCTTCAACAAGCTGCATCTGGACTTCCTCCTCGGTCGC
GGGGTCATGCTCCTGGGGCTCTGGAAGTTGTTCGACGTAGACCTGAAGGAGGAGCCAGCG
P  Q  Y  E  D  P  E  T  F  N  K  L  H  L  D  F  L  L  G  R

      5590                 5610                 5630
AGCTGACACAGACCCCGGCCGGTGCCGCCAACCCCTGCAACCCGGGCGGCACCGGCCGGA
TCGACTGTGTCTGGGGCCGGCCACGGCGGTTGGGGACGTTGGGCCCGCCGTGGCCGGCCT
S  *

      5650                 5670                 5690
TCTCACTTACCCGACCTATTGCGCTCTCGTCCGGACCCCCGGAGAGAAAGCGCCGAAGCA
AGAGTGAATGGGCTGGATAACGCGAGAGCAGGCCTGGGGGCCTCTCTTTCGCGGCTTCGT

      5710                 5730                 5750
GCAGCAAGGAGACCGCCGCGATGCCTGTAGCGCTGTGCGCGATGTCGCACTCCCCCCTGA
```

Fig. 4J

```
CGTCGTTCCTCTGGCGGCGCTACGGACATCGCGACACGCGCTACAGCGTGAGGGGGGACT
                    M  P  V  A  L  C  A  M  S  H  S  P  L  M

      5770                5790                5810
       .                   .                   .                   .
TGGGACGCAACGACCCCGAACAGGAAGTCATCGACGCCGTCGACGCCGCATTCGACCACG
ACCCTGCGTTGCTGGGGCTTGTCCTTCAGTAGCTGCGGCAGCTGCGGCGTAAGCTGGTGC
  G  R  N  D  P  E  Q  E  V  I  D  A  V  D  A  A  F  D  H  A

      5830                5850                5870
       .                   .                   .                   .
CGCGCCGGTTCGTCGCCGACTTCGCCCCCGATCTCATCGTCATCTTCGCCCCCGACCACT
GCGCGGCCAAGCAGCGGCTGAAGCGGGGGCTAGAGTAGCAGTAGAAGCGGGGGCTGGTGA
  R  R  F  V  A  D  F  A  P  D  L  I  V  I  F  A  P  D  H  Y

      5890                5910                5930
       .                   .                   .                   .
ACAACGGCGTCTTCTACGACCTGCTGCCGCCGTTCTGTATCGGTGCCGCCGCGCAGTCCG
TGTTGCCGCAGAAGATGCTGGACGACGGCGGCAAGACATAGCCACGGCGGCGCGTCAGGC
  N  G  V  F  Y  D  L  L  P  P  F  C  I  G  A  A  A  Q  S  V

      5950                5970                5990
       .                   .                   .                   .
TCGGCGACTACGGCACCGAAGCCGGCCCTCTCGACGTCGACCGTGACGCCGCCTACGCAG
AGCCGCTGATGCCGTGGCTTCGGCCGGGAGAGCTGCAGCTGGCACTGCGGCGGATGCGTC
  G  D  Y  G  T  E  A  G  P  L  D  V  D  R  D  A  A  Y  A  V

      6010                6030                6050
       .                   .                   .                   .
TCGCCCGCGACGTCCTCGACAGCGGCATCGACGTCGCATTCTCCGAACGCATGCACGTCG
AGCGGGCGCTGCAGGAGCTGTCGCCGTAGCTGCAGCGTAAGAGGCTTGCGTACGTGCAGC
  A  R  D  V  L  D  S  G  I  D  V  A  F  S  E  R  M  H  V  D

      6070                6090                6110
       .                   .                   .                   .
ACCACGGATTCGCCCAAGCACTCCAATTGCTGGTCGGATCGATCACCGCCGTGCCGACCG
TGGTGCCTAAGCGGGTTCGTGAGGTTAACGACCAGCCTAGCTAGTGGCGGCACGGCTGGC
  H  G  F  A  Q  A  L  Q  L  L  V  G  S  I  T  A  V  P  T  V

      6130                6150                6170
       .                   .                   .                   .
TGCCGATCTTCATCAATTCGGTCGCCGAACCGCTCGGCCCGGTCAGCCGGGTACGGCTGC
ACGGCTAGAAGTAGTTAAGCCAGCGGCTTGGCGAGCCGGGCCAGTCGGCCCATGCCGACG
  P  I  F  I  N  S  V  A  E  P  L  G  P  V  S  R  V  R  L  L

      6190                6210                6230
       .                   .                   .                   .
TCGGCGAGGCGGTCGGGCGGGCCGCTGCCAAGCTGGACAAGCGTGTGCTGTTCGTCGGAT
AGCCGCTCCGCCAGCCCGCCCGGCGACGGTTCGACCTGTTCGCACACGACAAGCAGCCTA
  G  E  A  V  G  R  A  A  A  K  L  D  K  R  V  L  F  V  G  S

      6250                6270                6290
       .                   .                   .                   .
CCGGCGGCCTGTCCCACGACCCGCCGGTCCCGCAGTTCGCCACCGCGCCAGAGGAAGTGC
GGCCGCCGGACAGGGTGCTGGGCGGCCAGGGCGTCAAGCGGTGGCGCGGTCTCCTTCACG
  G  G  L  S  H  D  P  P  V  P  Q  F  A  T  A  P  E  E  V  R

      6310                6330                6350
       .                   .                   .                   .
GCGAGCGGTTGATCGACGGCCGCAATCCCAGTGCCGCCGAACGTGATGCCCGCGAACAGC
CGCTCGCCAACTAGCTGCCGGCGTTAGGGTCACGGCGGCTTGCACTACGGGCGCTTGTCG
  E  R  L  I  D  G  R  N  P  S  A  A  E  R  D  A  R  E  Q  R
```

Fig. 4K

```
      6370                6390                6410
GCGTCATCACCGCCGGGCGGGACTTCGCCGCCGGCACCGCCGCCATCCAGCCACTGAACC
CGCAGTAGTGGCGGCCCGCCCTGAAGCGGCGGCCGTGGCGGCGGTAGGTCGGTGACTTGG
  V  I  T  A  G  R  D  F  A  A  G  T  A  A  I  Q  P  L  N  P

      6430                6450                6470
CCGAATGGGACCGGCACCTGCTCGACGTCCTCGCCTCCGGCGACCTCGAGCAGATCGACG
GGCTTACCCTGGCCGTGGACGAGCTGCAGGAGCGGAGGCCGCTGGAGCTCGTCTAGCTGC
  E  W  D  R  H  L  L  D  V  L  A  S  G  D  L  E  Q  I  D  A

      6490                6510                6530
CGTGGACCAACGACTGGTTCGTCGAACAGGCCGGACACTCCTCCCACGAAGTGCGCACCT
GCACCTGGTTGCTGACCAAGCAGCTTGTCCGGCCTGTGAGGAGGGTGCTTCACGCGTGGA
  W  T  N  D  W  F  V  E  Q  A  G  H  S  S  H  E  V  R  T  W

      6550                6570                6590
GGATCGCCGCGTACGCGGCAATGAGCGCCGCCGGGAAGTACCGCGTCACCTCGACCTTCT
CCTAGCGGCGCATGCGCCGTTACTCGCGGCGGCCCTTCATGGCGCAGTGGAGCTGGAAGA
  I  A  A  Y  A  A  M  S  A  A  G  K  Y  R  V  T  S  T  F  Y

      6610                6630                6650
ACCGCGAAATCCACGAGTGGATAGCAGGATTCGGGATTACTACCGCCGTCGCCGTCGACG
TGGCGCTTTAGGTGCTCACCTATCGTCCTAAGCCCTAATGATGGCGGCAGCGGCAGCTGC
  R  E  I  H  E  W  I  A  G  F  G  I  T  T  A  V  A  V  D  E

      6670                6690                6710
AATAGACCCCGCCGCTCCCGCCCCGCAGTCCCAACGAAGGGTGGCCCCGGATGACCTCCG
TTATCTGGGGCGGCGAGGGCGGGGCGTCAGGGTTGCTTCCCACCGGGGCCTACTGGAGGC
  *                                                   M  T  S  V

      6730                6750                6770
TCCGCCCGTGCTCGCCGTCGGTGAACGCGGGCTGGTCGGTGGGCAGGAAGACCTCATCGC
AGGCGGGCACGAGCGGCAGCCACTTGCGCCCGACCAGCCACCCGTCCTTCTGGAGTAGCG
  R  P  C  S  P  S  V  N  A  G  W  S  V  G  R  K  T  S  S  P

      6790                6810                6830
CGACATCGCCCTCGACCTCGCAGCTCGTCAGTAGGAATGCGCACGGGCCGACGAGTCGCG
GCTGTAGCGGGAGCTGGAGCGTCGAGCAGTCATCCTTACGCGTGCCCGGCTGCTCAGCGC
  T  S  P  S  T  S  Q  L  V  S  R  N  A  H  G  P  T  S  R  A

      6850                6870                6890
CTGGTCACCGGGGCCAGCCGCGGCATCGGGGCGGCCATCGCAGATGCGGTGGCCGCCTCC
GACCAGTGGCCCCGGTCGGCGCCGTAGCCCCGCCGGTAGCGTCTACGCCACCGGCGGAGG
  G  H  R  G  Q  P  R  H  R  G  G  H  R  R  C  G  G  R  L  R

      6910                6930                6950
GGTGCCGCCGTAATCGTCCACTACGGATCCGATCGGACGGCCGCCGCTGCGGTGTCGACG
CCACGGCGGCATTAGCAGGTGATGCCTAGGCTAGCCTGCCGGCGGCGACGCCACAGCTGC
  C  R  R  N  R  P  L  R  I  R  S  D  G  R  R  C  G  V  D  G

      6970                6990                7010
GCATCACGGCTGCCGGGGGCCTCGCGGCTGCGGTCCAGGCCGACCTGTCCCGACCCGAGG
```

Fig. 4L

```
CGTAGTGCCGACGGCCCCCGGAGCGCCGACGCCAGGTCCGGCTGGACAGGGCTGGGCTCC
  I  T  A  A  G  G  L  A  A  A  V  Q  A  D  L  S  R  P  E  G

      7030                 7050                 7070
      .    .    .    .    .    .    .    .    .    .    .    .
GGCCTGAAGAGCTGATGCGGGAGTTCGACTCCGCGCTCGACGGTCTCGGGCTCGACCGAG
CCGGACTTCTCGACTACGCCCTCAAGCTGAGGCGCGAGCTGCCAGAGCCCGAGCTGGCTC
  P  E  E  L  M  R  E  F  D  S  A  L  D  G  L  G  L  D  R  G

      7090                 7110                 7130
      .    .    .    .    .    .    .    .    .    .    .    .
GGCTCGACATCCTCGTCAACAACGCCGGAATCAGTCGGCGCGGAGCGCTCGAGCGCGTCA
CCGAGCTGTAGGAGCAGTTGTTGCGGCCTTAGTCAGCCGCGCCTCGCGAGCTCGCGCAGT
  L  D  I  L  V  N  N  A  G  I  S  R  R  G  A  L  E  R  V  T

      7150                 7170                 7190
      .    .    .    .    .    .    .    .    .    .    .    .
CTGTCGAGGATTTCGACCGTCTGGTCGCACTCAACCAGCGCGCCCCGTTCTTCGTGACTC
GACAGCTCCTAAAGCTGGCAGACCAGCGTGAGTTGGTCGCGCGGGGCAAGAAGCACTGAG
  V  E  D  F  D  R  L  V  A  L  N  Q  R  A  P  F  F  V  T  R

      7210                 7230                 7250
      .    .    .    .    .    .    .    .    .    .    .    .
GGCATGCCCTGCCCCGGATGCACGACGGCGGTCGCATCGTCAACATTTCCTCCGGATCCG
CCGTACGGGACGGGGCCTACGTGCTGCCGCCAGCGTAGCAGTTGTAAAGGAGGCCTAGGC
  H  A  L  P  R  M  H  D  G  G  R  I  V  N  I  S  S  G  S  A

      7270                 7290                 7310
      .    .    .    .    .    .    .    .    .    .    .    .
CCCGCTACGCCAGACCCGACGTCATCAGCTACGCCATGACCAAGGGGGCGATCGAGGTGC
GGGCGATGCGGTCTGGGCTGCAGTAGTCGATGCGGTACTGGTTCCCCCGCTAGCTCCACG
  R  Y  A  R  P  D  V  I  S  Y  A  M  T  K  G  A  I  E  V  L

      7330                 7350                 7370
      .    .    .    .    .    .    .    .    .    .    .    .
TCACCCGCGCCCTCGCCGTAGACGTCGGCGAACGAGGCATCACCGCCAACGCCGTGGCGC
AGTGGGCGCGGGAGCGGCATCTGCAGCCGCTTGCTCCGTAGTGGCGGTTGCGGCACCGCG
  T  R  A  L  A  V  D  V  G  E  R  G  I  T  A  N  A  V  A  P

      7390                 7410                 7430
      .    .    .    .    .    .    .    .    .    .    .    .
CGGCCGCGCTCGATACCGACATGAACGCGCACTGGCTTCGCGGTGACGACCATGCCCGCA
GCCGGCGCGAGCTATGGCTGTACTTGCGCGTGACCGAAGCGCCACTGCTGGTACGGGCGT
  A  A  L  D  T  D  M  N  A  H  W  L  R  G  D  D  H  A  R  T

      7450                 7470                 7490
      .    .    .    .    .    .    .    .    .    .    .    .
CCACCGCCGCGTCCACCACTGCACTGCGAAAACTCGCCACCGCGGAGGACATCGCCGCGA
GGTGGCGGCGCAGGTGGTGACGTGACGCTTTTGAGCGGTGGCGCCTCCTGTAGCGGCGCT
  T  A  A  S  T  T  A  L  R  K  L  A  T  A  E  D  I  A  A  I

      7510                 7530                 7550
      .    .    .    .    .    .    .    .    .    .    .    .
TCGTGGCCTTCCTCGTCAGCGCCGCCGCCGGTGCGATCACCGGGCAGGTCATCGACGCCA
AGCACCGGAAGGAGCAGTCGCGGCGGCGGCCACGCTAGTGGCCCGTCCAGTAGCTGCGGT
  V  A  F  L  V  S  A  A  A  G  A  I  T  G  Q  V  I  D  A  T

      7570
      .    .
CCAACGGCAACCGGCTCTAACCAG
GGTTGCCGTTGGCCGAGATTGGTC
  N  G  N  R  L  *
```

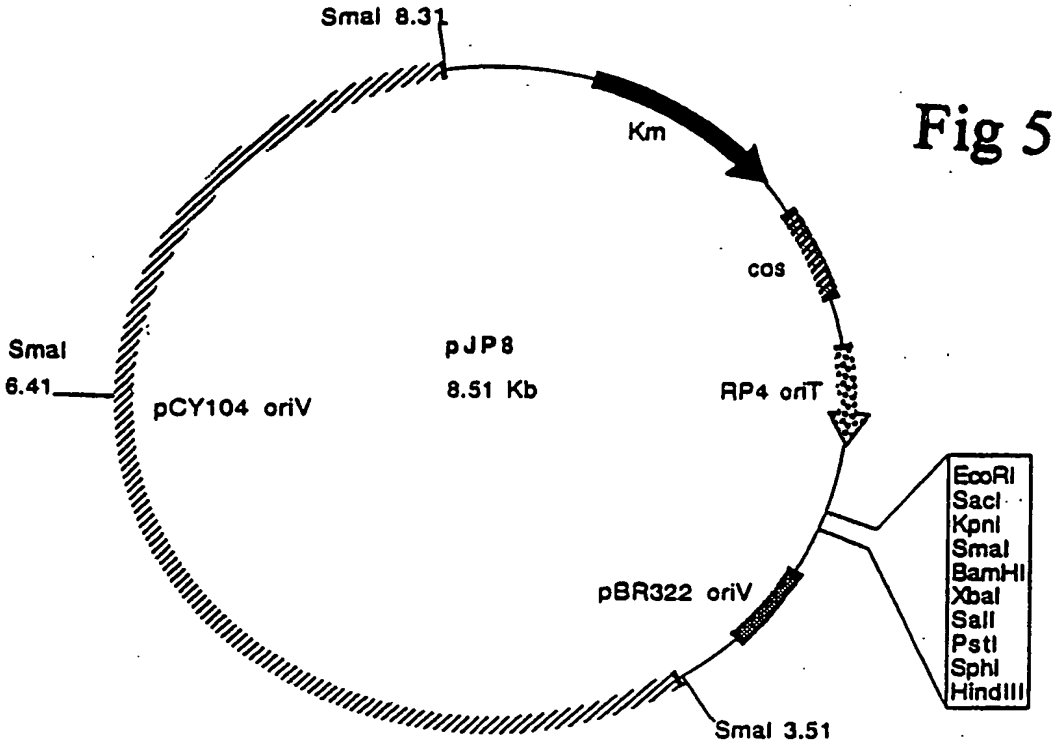


Fig 5

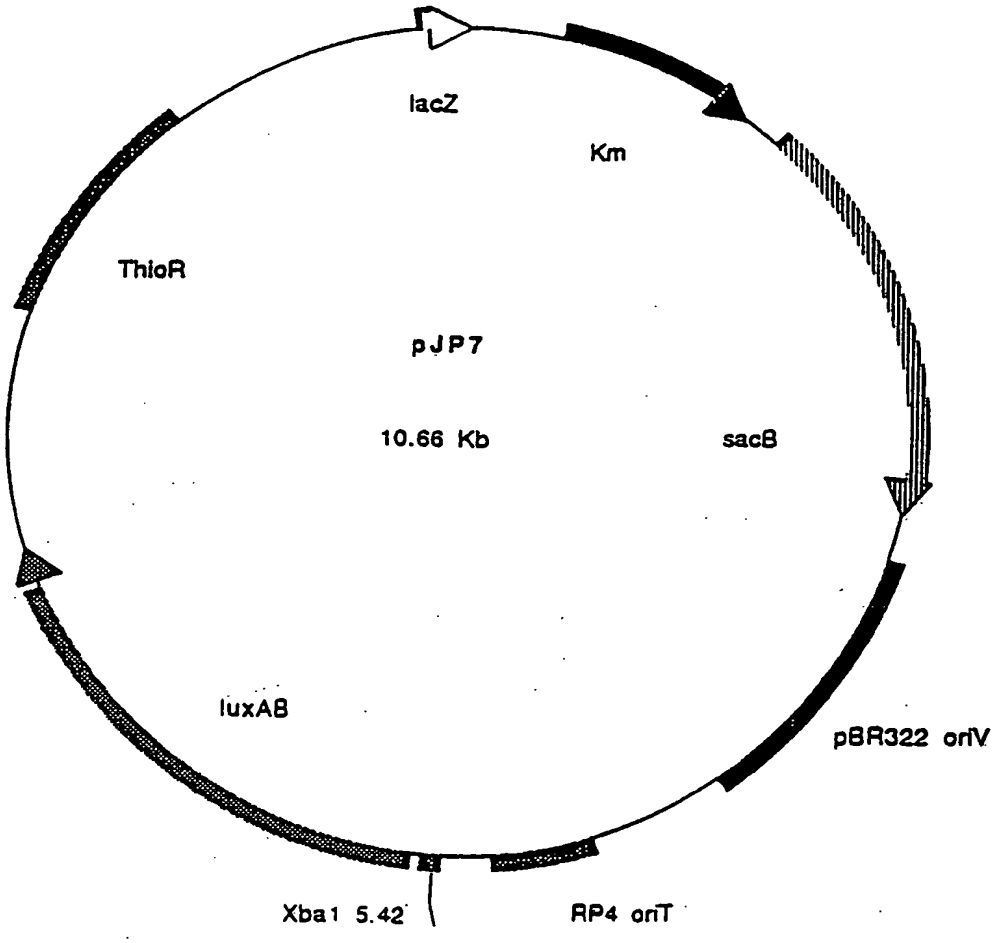


Fig 6

Fig 7

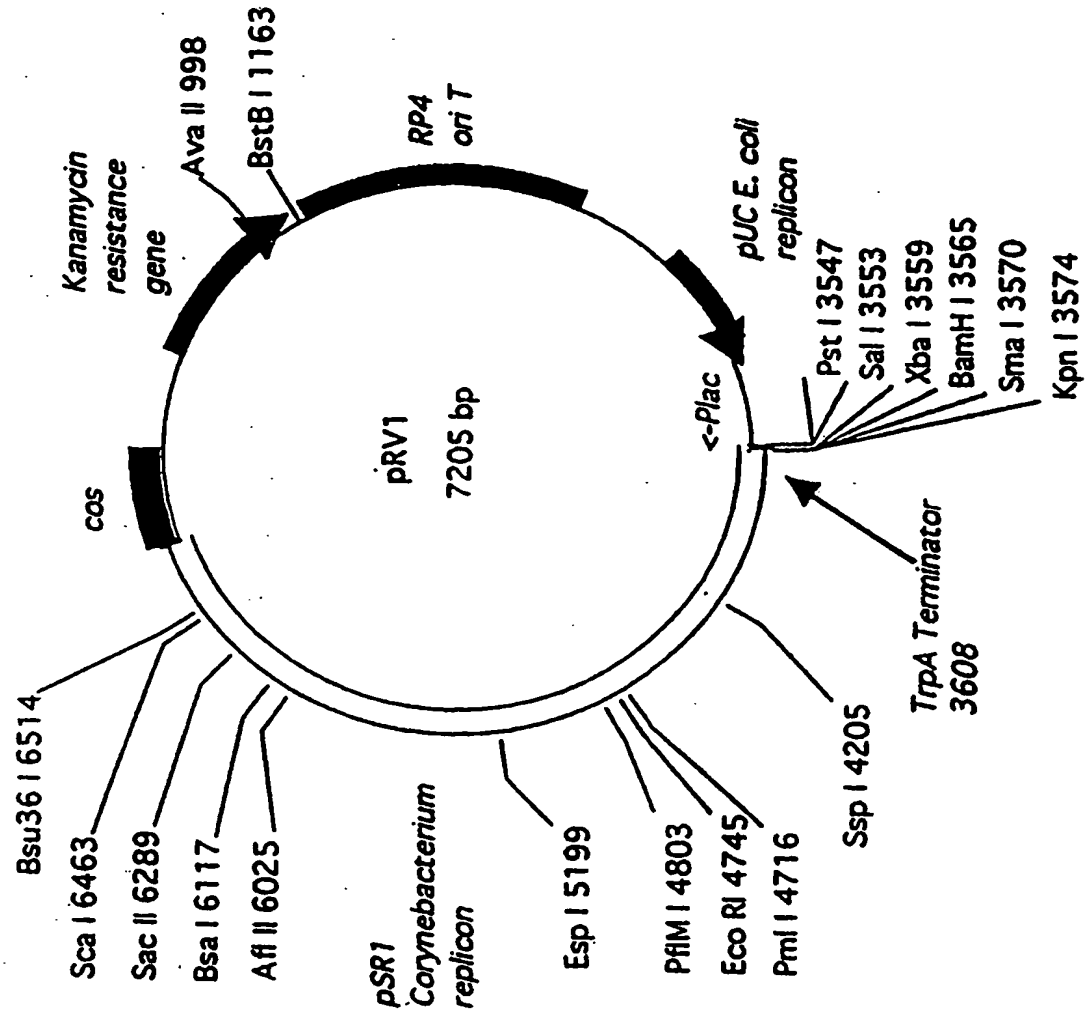


Fig 8